THE NATURE AND STRUCTURE OF SCIENTIFIC THEORIES

C. U. Moulines
Seminar für Philosophie, Logik und Wissenschaftstheorie,University of Munich, Germany

Keywords: axiom system, axiomatization, basic concept, constructive empiricism, data model, definition, empirical claim (of a theory), instrumentalism, metatheoretical structuralism, model, modelistic conceptions (of theories), observational concept, radical empiricism, scientific realism, set-theoretical view (of theories), theorem, theoretical concept, two-levels view (of theories), truth (of a theory).

Contents

UNESCO – EOLSS SAMPLE CHAPTERS

Summary

This chapter is devoted to two central questions in philosophy of science that are closely related to each other: What are the adequate criteria to identify scientific theories (to distinguish one theory from another)? How are theories scientifically (ideally) structured? It is argued that, to answer these questions, a first step to be made is to discuss the issue of the determination of the meaning of scientific concepts. For this, the axiomatic method to lay out scientific theories in a completely precise way is of great value. However, in the case of empirical theories, this method alone is insufficient. A consideration of how scientific concepts and principles are ‘anchored’ in experience is required too. The (semantic) notion of a model proves to be crucial here. We examine several epistemological conceptions that have been proposed to deal with this problem: radical empiricism, the ‘two-levels’ conception, the set-theoretical view of theories, constructive empiricism, and metatheoretical structuralism. They provide an increasingly sophisticated picture of the identity and structure of scientific theories. Finally, the issue of the truth of scientific theories will be discussed briefly.

1. Introduction

Empirical science is a complex building consisting of many different components: instruments and methods of observation, experimentation and computation, technological applications, methodological and ethical values, underlying ideological and/or metaphysical motivations and assumptions, and scientific communities studying a particular range of human experience with some particular goals in mind. But above all, empirical science consists of a particular sort of abstract entities known as *theories*. Instruments, methods, values, goals, research communities and all the rest make sense only with respect to some particular theories accepted and used by scientists. The notion of a scientific theory is essential to understand the nature of empirical science. Therefore, it is very important for the philosophy of science to make clear what kind of entity a (scientific) theory is, how it is articulated, and how it works. The analysis of the identity criteria, structure and functioning of scientific theories in general is the topic of this chapter. We shall examine the main systematic approaches that have been proposed for this kind of analysis in the philosophy of science of the last hundred years or so, and compare their relative merits and shortcomings.

In the present chapter, the use of the term "(scientific) theory" will essentially be restricted to theories constructed within the context of *empirical* science, i.e. theories that, in the last analysis, directly or indirectly, have some link to human (sensorial) experience. The consideration of purely logical or mathematical theories, or of metaphysical or ideological theories, falls out of the scope of this chapter. Nevertheless, some of the insights obtained in the philosophy of logic and mathematics are relevant to some issues concerning empirical theories as well, and some of the metatheoretical claims made about the nature of empirical theories also apply to purely logico-mathematical theories.

In present-day philosophy of science it is generally agreed that the idea of an 'absolute experience', completely independent of any theoretical considerations, is untenable, at least in science. For this reason, two questions become central in the discussion of the nature of empirical science: 1) What is a (scientific) theory, i.e. what is the 'essence' of a scientific theory, how is it built up, how does it work? And: 2) How does a theory relate to its corresponding experiential basis? The two questions are obviously interrelated. To deal with these two questions modern philosophy of science has devised various 'models' on the nature and working of scientific theories - i.e. it has devised its own *metatheories*, as it were, on theories. Some of these models, or rather some aspects of these models, are widely held within the community of philosophers of science, but others are still being discussed quite controversially.

In very general terms, the first thing to notice about scientific theories is that they are particular cultural items devised by some people (usually called "scientists") at some particular time in the cultural history of mankind with a particular aim in mind: to systematize, explain and predict some phenomena we are confronted with in a particular area of our experience, and to do so in a methodologically well-founded manner, and not just by speculation. In short, scientific theories are products of the human mind devised to make well-founded claims about the world, or, to put it in other words, to set up well-founded propositions with an empirical content. Now to make any claims about

UNESCO – EOLSS SAMPLE CHAPTERS

anything we need to have previously, as a matter of course, a set of *concepts* by means of which the claims or propositions are formulated. These concepts are expressed as *terms* of a particular language, be it a natural language like English, or a purely mathematical language like the differential calculus or, more often, a mixture of a natural and a mathematical language. Most classical authors in philosophy of science would contend, explicitly or implicitly, that a theory is nothing but a 'thing' consisting simply of a set of concepts and a set of propositions constructed by means of the concepts, or in a more linguistic way of speaking, consisting of terms of a given language and statements constructed with those terms. More recent approaches in the philosophy of science contest this view of theories, which is considered to be too simple-minded, and would rather take at least the notion of a model (which is not a linguistic entity) as an essential constituent of a theory. We shall come back to this controversial issue below. At any rate, there is a wide consensus that concepts (or, linguistically speaking, terms) are the most basic units in the structure of a any given theory. (In this chapter, no clear-cut distinction will be made between "concepts" and "terms": We'll speak of terms when a particular language is presupposed, and of "concepts" in a more general setting.) Consequently, a first general question that arises in the metatheory of scientific theories is the question about the nature of scientific concepts, and more particularly, the questions of how the meaning of scientific concepts is determined and how they are put together in order to make statements or propositions about the world.

2. The Problem of the Meaning of Scientific Concepts

2.1. The Semantic Specificity of Scientific Concepts

The first important step when constructing a scientific theory consists in the indication of a series of *specific concepts*. Any particular scientific theory aims at the investigation of a particular domain of our experience, but it does so by assuming a specific, thoroughly worked-out conceptual framework. This means that the domain of experience to be investigated has to be, first of all, 'interpreted' or 'reconstructed' in terms of the assumed conceptual framework. As has already been pointed out, the idea of a 'pre-conceptual' experience is untenable in a scientific context. Consider, for example, the domain of experience relevant for the science of mechanics: This domain is interpreted or conceived in terms of such concepts as *particle, position, time, velocity, mass,* etc. On the other hand, if we want to deal with the domain corresponding to the aims of decision theory, we will use such notions as *action, uncertainty, expected utility, subjective probability,* and the like. Clearly, the second group of concepts is very different from the first one. Each of these groups corresponds to a specific conceptual framework within which a particular theory about a particular empirical domain has to be laid out. When dealing with the motion of medium-sized material objects we employ the first group of concepts; when dealing with decisions that have to be taken by humans beings, we may use the second group.

It is important to notice the specificity and precision of the concepts relevant for the construction of an adequate scientific theory. They are quite different in nature from the usual notions we employ when accounting for everyday experience. True, scientific concepts are often *expressed* by means of words coming from everyday language (a

UNESCO – EOLSS SAMPLE CHAPTERS

feature that usually suggests the historical origins of the discipline in question); however, these expressions normally have a heavily transformed, and above all more precise, usage when compared with everyday language. The usage of the English words "force" or "field" in a physics textbook is only remotely, if at all, related to their usage in everyday English. Moreover, in many scientific theories we find terms that are complete neologisms (think of "entropy", "spin", "gene", ...) and that either have no use in everyday language, or else, if they now have such a use, it has been imported from science ("gene" would be a good example of this). The rationale for introducing such terms is not the wish to disorient the layman and invent a cryptic language only known to a clique, but rather to avoid false associations and expectations, and to be as precise as possible.

By stressing the deep semantic differences between genuinely scientific concepts and everyday concepts we don't want to suggest, however, that there is an absolute cut dividing both kinds of concepts. In some contexts, it may be convenient to establish significant connections between some specifically scientific concepts and some everyday concepts. Moreover, in some disciplines, especially when they are at the beginning of their development, the concepts used are more or less taken directly from everyday experience, and they change their meaning gradually with the development of the theory. The more a discipline approaches its mature status, the more the differences in the semantic structure of its concepts with respect to their everyday origins becomes striking. In a word, the distinction between scientific and everyday concepts is a gradual and historical one, but it is nevertheless a very significant one for the analysis of the semantic structure of scientific theories.

The semantic specificity of scientific concepts poses a fundamental problem: Their meaning cannot be settled or grasped in the same way the meaning of everyday concepts is. As small children, we all learn quite successfully to get along with such concepts as "red", "warm", "dog", "table", and so on, without need of much reflection or 'theory'. It suffices to become acquainted with familiar experiences and situations. In the case of specifically scientific concepts, the everyday way of grasping concepts is not possible any more, or at most only to a highly limited extent.

Now, if the meaning of scientific terms usually doesn't coincide with the meaning of everyday expressions, the question immediately arises: How do they obtain their proper meaning? This is no trivial question at all; rather, it reflects a quite central problem in the philosophy of science. The answer to this question is all but simple. The different approaches that have been offered to answer it belong, at least in part, to some of the most controversial points of modern philosophy of science. We'll come back to this question below. For the moment, let's just make the following remark. Whatever the case may be for everyday concepts, it is for sure that scientific concepts never obtain their proper meaning one by one - there is no such thing as an 'isolated' scientific concept; they get their meaning in the context of a whole series of *other well-determined concepts*; together, they build a specific conceptual framework.

2.2 The Definition of Scientific Concepts: Its Possibilities and Limits

It is often said that the standard way to determine the meaning of scientific concepts is,

UNESCO – EOLSS SAMPLE CHAPTERS

other than in the case of everyday notions, to *define* them rigorously. This is only partially true, and this for two reasons. First, to define a term A means to set it in a systematic relationship with other terms, say, $B, C, D, \ldots$. Now, if this is going to be a real definition, it has to be guaranteed that the information content provided by A is exactly the same as the one provided by the combination of $B, C, D, \ldots$. A precise way to express this situation is to say that A and the combination of $B, C, D, \ldots$ have to be *semantically equivalent*. (For example, the notion of average velocity in mechanics may be defined by means of the concepts "distance" and "time interval" because the first notion is semantically equivalent to "distance run divided by elapsed time interval".) However, not all conceptual connections appearing in a scientific context can be interpreted as such *semantic equivalences*. In many cases, what we have are only more or less partial connections that don't amount to a full conceptual coincidence. (Think of 'defining' the energy of a gas, in thermodynamics, as the partial derivative of pressure with respect to volume: This can only be accepted if other relevant parameters for the state of the gas are supposed to be held constant.)

The second reason why definitions cannot be the general rule for determining the meaning of scientific concepts is more fundamental. It is just logically impossible to define *all* relevant notions of a given scientific discipline: This would lead either to an infinite chain of definitions (an absurd idea), or else to a vicious circle. (Suppose, to simplify, that our discipline would contain only three concepts, A, B, C ; we should first define A in terms of B and C ; but then we should define B in terms of A and C - therefore not defining anything at all; the argument is applicable to any finite set of concepts.)

The consequence of this logical fact is that, in any given theory, we have to admit a certain number of concepts as *undefined*. In the terminology usual in formal philosophy of science, such notions are called "primitive concepts" or also "basic concepts". (In this chapter, we'll prefer the latter terminology.) Once we have admitted one such set of basic concepts, we have to take care that the rest of the theory's specific notions can be introduced as defined concepts by means of rigorously constructed definition chains, all of its members being semantic equivalences. In terms of formal logic, such semantic equivalences are usually formulated as generalized *biconditionals* (stated by means of the logical connective "if and only if"). Suppose, for example, that we start with three basic concepts, expressed by the simple monadic predicates "$A(x)$", "$B(x)$", and "$C(x)$"; then we can introduce a first defined predicate "$D(x)$", for example, by means of the biconditional

Def. 1: For all x, $D(x)$ if and only if $A(x)$ and $B(x)$

Then, we could introduce a further defined concept "E" by means of, say, the biconditional

Def. 2: For all x, $E(x)$ if and only if $A(x)$ and not-$B(x)$

Finally, we could still introduce a further defined concept "F" by means of

Def. 3: For all x, $F(x)$ if and only if $E(x)$ or $C(x)$

We can then say that there is a *definition chain* leading from the basic concepts "A", "B", "C" to the defined concept "F", consisting of the three definitions Def. 1 – 3.

Since the pioneer times of formal philosophy of science there is wide consensus about the strict rules that definitions and definition chains have to fulfill in order to be adequate. We cannot go into the explication of these rules here. Suffice to note that their role is to guarantee that, among other things, the definitions are not circular, not contradictory, and not 'creative' (i.e. they are really semantic equivalences and not concealed statements of a matter of fact.) At any rate, it is quite clear how defined concepts get their meaning in the conceptual framework of a theory: through definition chains that eventually lead back to the meaning of the basic concepts. But now we are still confronted with the fundamental question of how the basic concepts get their meaning determined at all. How is the meaning of the concepts "A", "B", and "C" in our example determined if they cannot be led back to other, previously given concepts?

3. The Axiomatic Construction of a Scientific Theory

3.1. The General Idea

To deal with the last question, we have to deepen our analysis of the structure of scientific theories and get thereby to the core of the question of the structure of such theories. It has already been pointed out that scientific concepts don't appear in isolation but in 'clusters'. This fact is particularly significant for the problem of determining the meaning of the basic concepts of a given theory. Typically, such concepts don't appear isolated in simple statements; rather, they appear inter-connected in the theory's general, fundamental principles. In the previous section we have seen that a theory consists, in a first move, of a specific conceptual framework; but the conceptual framework is only one of the components of the identity of any theory. It is a framework within which some statements of fact about the world are supposed to be made; and only when these statements appear to be justified, can we assume that we have obtained some knowledge about the world – which is, after all, the main purpose scientific theories aim at. Certainly, the definitions by means of which new concepts are introduced into the conceptual framework on the basis of previous ones, are also statements since they are assertions about the semantic equivalence of certain concepts. However, definitions alone don't provide any substantial knowledge about the world. For sure, in a given theory we could assert, for example, that the term "babaloo" is semantically equivalent to the combination of the basic terms "abracadabra" and "tillybilly"; but so long as we don't know what "abracadabra" and "tillybilly" refer to, we haven't gained any substantial knowledge.

Now, the most significant statements a theory makes are precisely those that are essential to obtain the knowledge we aim at, and they cannot be just definitions. They are the theory's *axioms*. In principle, axioms are the assertions we make by means of the *basic* concepts. For reasons of convenience, sometimes assertions that establish connections between *defined* concepts are also introduced as axioms; this may make the exposition of the theory more perspicuous or easier to understand; however, even in

UNESCO – EOLSS SAMPLE CHAPTERS

such cases, the axioms formulated with the help of defined concepts should be in principle translatable into axioms formulated only with basic concepts. The point is that, at least in principle, all other statements of fact we want to make about the world within the theory's conceptual framework have to be derived logically from the axioms (frequently together with some definitions) as *theorems*. The fundamental semantic question about the meaning of the basic concepts is thus tied to their occurring in the axioms of a theory and to the way in which the axiomatic statements where the basic concepts appear tell us something about the world. We'll come back to this issue later on. For the moment, let's examine in some more detail what the setting up of axioms in a given domain of our knowledge implies for understanding the essential structure of a scientific theory.

Axioms, definitions, and theorems are the three fundamental categories of statements building a scientific theory. As we have seen, definitions don't provide genuine knowledge; their purpose is rather the practical aim of obtaining more easily, or 'elegantly', the knowledge expressed in the theorems when deducing them from the axioms. And all what is stated in the theorems, since they are just deductive consequences of the axioms and definitions, is already implicitly contained in the axioms. Thus, in the last analysis, anything a theory asserts about the world is already contained, explicitly or implicitly, in the axioms.

Whenever a theory is constructed in such a rigorous way that we may clearly determine which statements are genuine definitions, which ones are axioms, and which ones are theorems, then we say that the theory has been *axiomatized*. And the set of rules and procedures which lead to the axiomatization of any given theory is called "*the axiomatic method*".

In older presentations of the axiomatic method, a further distinction was made between two kinds of basic principles of a theory: genuine axioms and so-called "postulates". The first kind was supposed to be common to all scientific theories, or at least to a great number of them, and were usually thought of as logical or metaphysical ('synthetic a priori') principles, while the postulates were supposed to be the basic statements valid only in a particular theory. Today, this distinction has become rather obsolete, since it is not at all clear that there really are principles assumed in *all* possible scientific theories. (Even logic is not any more regarded as so universal as it was previously thought to be.) Consequently, most authors today make no essential distinction between the terms "axiom" and "postulate", and treat them as synonymous. In what follows we we'll restrict our usage to the term "axiom".

The idea that axiomatization is essential to make the genuine structure of a scientific theory explicit and, that it therefore settles the theory's identity completely, has a long history. It goes back at least to Aristotle. Probably inspired by Aristotle's general ideas, shortly afterwards Euclid brought all geometrical knowledge that had been gathered until then into an axiomatic form. The result was what we now know as "Euclidean geometry". Historically, this represents the first concrete example of a theory presented in axiomatic form – an example that much impressed one generation after another of scientists and philosophers. However, it was not the only important example of axiomatization of a theory in Ancient times: Roughly a century later, Archimedes

UNESCO – EOLSS SAMPLE CHAPTERS

axiomatized the theory of statics. And some other minor attempts at axiomatization of different disciplines were made in the Hellenistic times. In the 17th century, Descartes and Newton tried to axiomatize what we now know as “classical mechanics”. Newton was more successful than Descartes in this endeavor but, still, his presentation didn’t really match the standards that had been set by Aristotle, Euclid, and Archimedes. (The reason for this partial failure is not that Newton was a worse scientist than his Greek predecessors but rather that the knowledge implicitly contained in general mechanics is much more complex than geometrical knowledge or the knowledge about static relationships.) Starting by the end of the 19th century, huge and quite successful efforts were made all along the 20th century to axiomatize a great number of theories of all scientific disciplines: Not only almost all significant mathematical theories, but also many important theories of the empirical sciences – classical mechanics, thermodynamics, quantum mechanics, relativistic space-time theory, genetics, and a number of theories of the social sciences – have been axiomatized in a quite satisfactory way. For this, the use of the formalization tools provided by modern logic since the end of the 19th century has been quite instrumental.

It has to be remarked, however, that the axiomatization of a theory very often represents rather an ideal (a ‘regulative principle’) and not so much a reality within scientific practice. Only within the formal disciplines of logic and mathematics may be said that almost all existing theories have been completely and rigorously axiomatized. In the empirical sciences, the axiomatic construction of a theory is a much more involved and problematic matter. Even in the case of the rather successful examples of axiomatizations of empirical theories we have just mentioned, a cursory look at them already makes clear that they cannot represent the whole content of the theory as intuitively grasped before the axiomatization has taken place. That is, the result of the axiomatization, even in the cases where it has been carried through most conscientiously, always appears essentially incomplete in the sense that statements that intuitively should be regarded as belonging to the theory in question cannot be derived from the set of axioms proposed. This situation is to some extent due to historical contingencies (the axiomatic tradition being much stronger in mathematics than in empirical science since Antiquity). But the more fundamental reason for the irremediable incompleteness of the best axiomatizations of empirical theories comes from a profound epistemological and methodological problem: the involved nature of the relationship between the theory and ‘the world outside there’ the theory is supposed to deal with – a problem logico-mathematical theories don’t have since they are, by their very nature, self-contained. As we shall see, the complexity of this relationship makes the structure of empirical theories also substantially more complex than in the case of purely mathematical theories; and this, in turn, as we shall see, has to do with the vexing problem already discussed at the beginning of this chapter of the meaning of the basic concepts of a theory. The identity criteria for empirical theories and the way their basic concepts are semantically settled are, though not completely, nevertheless quite substantially different from those corresponding to purely mathematical theories. To address this problem, different approaches to the understanding of the identity and structure of empirical theories have been proposed in the last one hundred years or so. They all have in common that they consider the axiomatic method quite important for analyzing the nature of empirical theories, but only as a *first step* in the endeavor. That is, they all consider that providing a list of basic concepts and of the axioms they occur

UNESCO – EOLSS SAMPLE CHAPTERS

in is very important in order to understand what the theory *is*, but that something more should be added if we want to have a complete picture of the nature of an empirical theory. In other words, in the case of empirical theories (other than in the case of logic and mathematics) axiomatization by itself is a necessary but not sufficient condition for identifying and analyzing them. Now, before we go on to examine the different approaches proposed to deal with this issue, let's consider a very simple example of axiomatization. This will help us in clarifying not only the notions discussed so far but it will also be useful as a concrete reference for the considerations later on.

3.2. A Simple Example of Axiomatization

We have said that the first step to get a clear picture of the identity of a theory is to lay it out in an axiomatic form. Now, it should always be borne in mind that scientific theories, especially in the case of empirical science, usually are not constructed from the scratch as a system of axioms. This is almost always the result of long and arduous efforts that may involve several generations of scientists (and philosophers). When Euclid axiomatized geometry, he had before his eyes a great amount of geometrical knowledge provided by previous mathematicians for at least two hundred years before him. Similarly, it might be said the Newton's axiomatization of mechanics not only summarized knowledge that Newton himself had obtained but also synthesized a century of strenuous research in mechanics. Even today, scientific textbooks hardly present their theories in an axiomatic form (even though the axiomatization might have already been provided in the specialized literature). So, the situation the axiomatic methodologist is confronted with is usually this. There is already a substantial amount of knowledge on a given area of our experience expressed in a number of specific concepts and assertions, which we *intuitively* regard as pertaining to one and the same theory; the task is then to find out which concepts should be regarded as basic so that all other concepts in that area can be defined through them, and which principles serve as axioms, so that all other statements found in the area in question can be reconstructed as mere definitions or else may be deduced as theorems from the chosen axioms and/or definitions. Of course, we could perform this task very easily by taking *all* concepts as basic concepts and *all* assertions made with them as axioms. From a purely formal point of view, nothing hinders us in doing so. But from an epistemological and methodological standpoint this procedure would be absurd, since our aim is to have a clear picture of the real content of the theory, and for this we have to settle the minimal set of concepts and principles that synthesizes that content. This is the task the 'axiomatizer' has to perform.

Consider the following example. (It is not an example of a 'serious' scientific theory but it will help us in clarifying important points about the axiomatic method and the structure of scientific theories.) Suppose in a given (let's say, sociopsychological) context, we find a mini-theory about family relationships, which tells us that a standard family is a group of people for which following propositions are asserted (the statements appear numbered for more convenience):

"(1) A person that has engendered another one is called his or her parent. (2) Nobody is engendered by just one parent. (3) Nobody is engendered by more than two parents. (4) The parents are married to each other. (5) Nobody is married to more than one person.

UNESCO – EOLSS SAMPLE CHAPTERS

(6) Persons engendered by the parents are called their children. (7) No child is married to one of his or her parents. (8) Two persons who are children of the same parents are called siblings. (9) If a person is a sibling of a second one, then the second is also a sibling of the first. (10) If a person is a sibling of a second one, and the latter is a sibling of a third one, then the first is also a sibling of the third. And so on. (The list could be continued indefinitely.)"

In addition to these explicitly asserted propositions about standard families, we should add some other propositions that very likely would not be asserted explicitly but are presupposed implicitly and would be asserted to by the theory's users if asked. Probably, they would be at least the following:

(11) Nobody can be engendered by himself or herself.
(12) Nobody can be married to himself or herself.
(13) If a person is married to another one, then the second is also married to the first.

(In more 'serious' theories, it will usually be a rather non-trivial task to make explicit propositions that are presupposed as a matter of course and are very important for the coherence and well-functioning of the theory but are not explicitly stated in its pre-axiomatic expositions.)

Call the set of statements (1) – (13) the "pre-axiomatic exposition of the theory of standard families", or "pre-axiomatic text" for short. Our task is now to bring it into a correct axiomatic form in order to make the conceptual and methodological structure of the theory of standard families (call it "TSF", for short) explicit. The two fundamental questions that have to be answered in order to fulfill this task are the following: What are the basic concepts of the theory? And what its axioms?

Take the concepts of *having engendered* and *being married to* as basic concepts. It is easy to see that all other concepts appearing in the theory can be defined by them. (We'll use the abbreviation "iff" for "if and only if".) For example,

Def. 1: For all x, y: x is a parent of y iff if x has engendered y.

Def. 2: For all x, y: x is a child of y iff y has engendered x.

(Or, alternatively)

Def. 2': For all x, y: x is a child of y iff y is a parent of x.

We leave to the reader the task of defining all other concepts appearing in the pre-axiomatic text by means of the two basic concepts chosen. Notice that different definitions may lead to the same defined concept. For example the concept of a child can be obtained either directly by Def. 2 or else by the chain Def. 1 – Def. 2'. Which chain we choose, is purely a matter of convenience, so long as we obtain a concept that has the same properties as stated in the pre-axiomatic text. Notice also (and this is a more important point) that we could have chosen a different set of basic concepts to define all other concepts appearing in the pre-axiomatic text. For example, instead of

UNESCO – EOLSS SAMPLE CHAPTERS

choosing "having engendered" as a basic concept, we could have chosen "being a parent of" as basic and then define "having engendered" as follows: " x has engendered y " iff " x is a parent of y ". All other concepts could then be obtained in the same way as before. This is an important feature of the axiomatic method in general, and though it seems to be rather trivial, it has remarkable consequences for the understanding of the identity of theories. We'll come back to this point in Section 7.4.

What about the axioms of our mini-theory of standard families? Take the following propositions (for more perspicuity we'll use the symbolic abbreviations " E " for "has engendered" and " M " for "is married to"):

Ax. 1: For all x, y: if $x\ E\ y$, then there is a z such that $z\ E\ y$, and $x \neq z$..

Ax.2:
For all x, y, z, w: if $x\ E\ w$, and $y\ E\ w$, and $z\ E\ w$, then either $x = y$, or $x = z$, or $y = z$.

Ax. 3: For all x, y: if $x \neq y$ and there is a z such that xEz and yEz, then xMy.

Ax. 4: For all x: if xMy, then yMx.

Ax. 5: For no x: xMx.

Ax. 6: For all x, y: if xMy, then there is no $z \neq y$, such that xMz.

Ax. 7: For all x, y: if xEy, then not-yEx.

It can be shown that all knowledge about standard families contained in the pre-axiomatic text is systematized and synthesized in these seven axioms (and the definitions). Some of the statements of the pre-axiomatic list would reappear now simply as definitions, like (1) and (6); others reappear as axioms, like (2), (5), (12), and (13); and still others are now theorems that can be deduced more or less directly from the axioms (and possibly the definitions). For example, statement (2) of the list is deducible from Def. 1 and Ax. 1; statement (3) follows from Def. 1 and Ax. 2, (4) follows from Def. 1 and Ax. 3, and (7) follows from Def. 1, Ax. 1, Ax. 2, Ax. 3, Ax. 6, and Ax. 7. Statement (8) could easily be reformulated as a new definition, from which (9) and (10) are deducible as theorems. Statement (11) is deducible from Ax. 7 as a theorem. (We leave the proofs to the reader.)

With this (admittedly highly simplified) example of an axiomatized theory in mind, let's now discuss the different views about the nature and structure of scientific theories that have been proposed in modern philosophy of science.

4. The Formalist Conception of Theories

Let's come back to the question we asked at the end of Section 2.2: How is the meaning of the basic of concepts of a scientific theory determined? In our simple example: How is the meaning of the terms " E " and " M " determined? Of course, we are supposed to

UNESCO – EOLSS SAMPLE CHAPTERS

know that "E" means what in ordinary English is usually meant by "having engendered" and "M" means what normal English speakers mean by "being married". But remember that, in a scientific context, we should mistrust the intuitive associations coming from everyday experience and the language we have learnt in our childhood. As we have remarked at the beginning of this chapter, at a certain stage of scientific development, at the latest, we should give up any intuitive, everyday associations that probably would be highly misleading. Suppose our mini-theory of standard families would have reached such a stage. Then, what about the meaning of "E" and "M"?

The answer of the so-called formalist philosophers of science, most prominently represented by David Hilbert in the first quarter of the 20th century, is that the meaning of the basic concepts of a theory is completely and exclusively settled by the axioms they occur in. In our example, this would mean that the meaning of "E" and "M" is totally settled by the axioms Ax. 1-7 – and nothing else. These axioms contain already all that has to be (scientifically) known about "having engendered" and "being married". To make this point clear, some authors introduced the expression "implicit definition": Axioms would be implicit definitions of the terms appearing in them. And the theory's identity would be exclusively given by the axioms with the terms 'implicitly defined' with their help. The theory of standard families, TSF, *is* just the set {Ax. 1, ..., Ax. 7}.

This conception has some degree of plausibility for theories of pure mathematics. After all, mathematical theories are self-contained in the sense that they don't pretend to claim anything about the world outside them – or, at least, so contends the formalist philosopher of mathematics. However, in the case of *empirical* science, this answer is clearly untenable, for the theories of empirical science are supposed to have, *by their very nature*, some empirical content, that is, some link to our experience of the real world. Their terms and axioms are intended to have some canonical interpretation in the world outside them. Providing some symbols as basic terms and some formulae as axioms where these terms appear is insufficient to characterize the theory: We have to add a further component to have a complete picture of its identity: an *interpretation*.

5. Theories as Interpreted Calculi

A system of formulae of a given language where some specific terms appear besides the notions of pure logic (such as variables, quantifiers, connectives and the symbol for identity of terms), whereby these terms occur in some formulae selected as axioms, and where all other statements are either definitions or may be deduced logically from the axioms, is called a "*calculus*". The system of statements {Ax. 1, ..., Ax. 7} as we have laid it out in Section 3.2 is an example of such a calculus. The view we examine in this section claims that scientific theories are a calculus *plus* a chosen interpretation (of the basic terms). Equivalently, we may say that a theory is an interpreted calculus. In our example, setting up {Ax. 1, ..., Ax. 7} would not be enough to characterize the theory of standard families: We have to interpret these axioms as stating some relationships between *real people*.

Now, of course, we have got to be more precise about what an interpretation actually is. For this, we need the formal notion of a *model* since the view we examine now claims

UNESCO – EOLSS SAMPLE CHAPTERS

that to interpret a calculus just means to provide it with a model (as a kind of structured entity). The model added to the calculus settles part of the meaning of the basic concepts – the part that was lacking in the purely formalistic view of theories. Notice that we are not saying that the formalist understanding of the meaning of basic concepts is completely wrong. It contains 'a grain of truth': The axioms provide a part of the meaning of the basic terms, while the model added to them provides the other part of the meaning. By appealing to the famous distinction introduced by Gottlob Frege between the sense and the reference of a term (both sense and reference being the essential components of its meaning), we might say that, in scientific theories, the axioms settle the sense of the basic terms while the model settles their reference.

The concept of a model we refer to here goes back to developments in formal semantics, especially in the work of Alfred Tarski. Claiming that a theory's conceptual framework can be interpreted in a particular domain of experience amounts to claiming that this domain (even though in a simplified or idealized form) can be conceived as a *model* of the theory's axioms. It is in this way that the concepts appearing in the axioms get their empirical content. In a different, though essentially equivalent, way of speaking, we may say that the theory (that is, its conceptual framework and the statements made within it) *represents* the domain in question by building a model. Still another way of putting it, is that a model is a structure constructed by means of the theory's concepts which *covers* the experiential domain we intend to study (in a more or less idealized manner).

We cannot go here into the precise definition of the notion of a model, which is a quite technical matter. Suffice to say that a model of an axiom system is a set (or a system of sets) of entities structured by some relations between them such that this (these) set(s) of entities and the relations between them make the assertions stated by the axioms *true*. An example may make this idea clearer. Take a group of people, say, the Hendersons, consisting of four persons (the entities of the model): John, Mary, Liz, and Pat. John and Mary are married, and they are the parents of Liz and Pat. Then, we can say that the Hendersons together with the relations just stated they have between them constitutes a model of the axioms of TSF because the statements Ax. 1, ..., Ax. 7 are all true of them, as can easily be verified. Note that, in order to arrive at this conclusion, there is a huge amount of details about the Hendersons we don't need to know: Neither their age or nationality, nor the place they live in, nor the colour of their hair, not even their gender, play a role. That is, to construct a model for TSF, we ignore a lot of information we could get about the Hendersons, which is just irrelevant to find out what TSF's axioms (and their basic concepts) might refer to. In this sense, the 'Hendersons-model' is a tremendously simplified or 'idealized' version of the 'real' Hendersons; but it is nevertheless a part of empirical reality, and this part is enough to establish that TSF has something to do with reality, that it is not just a game with symbols, not just a calculus but an *interpreted calculus*. This is a general feature of model construction in empirical science: The models we assume for our theories are usually quite 'abstract', 'ideal' constructs out of a much more complex reality, but they have nevertheless a real character by themselves and are not purely conceptual or verbal inventions.

The set of entities of a model which makes the axioms of a given theory true is called the "universe of discourse" of the theory. In our example, the Hendersons are the

UNESCO – EOLSS SAMPLE CHAPTERS

universe of discourse of TSF. We say, then, that the variables in the axioms of TSF "run over" the universe of discourse, in this case over the particular individuals in the Henderson family. We can also say that the variables $x, y, z, \ldots$ are interpreted over the set {John, Mary, Liz, Pat}, and that the symbol "E" appearing in TSF's axioms is interpreted as the relationship existing between John and Liz, John and Pat, Mary and Liz, and Mary and Pat, while the symbol "M" is interpreted as the relation between John and Mary. Note that in this way of depicting the relationship between the two components of a theory, the axiom system and its model, we assume, explicitly or implicitly, the construction of a systematic assignment of symbols of the axiom system (variables and basic terms) to real entities (say, real persons and real properties or relationships they have). Such an assignment is called an *interpretation function*. Strictly speaking, therefore, the conception of theories as interpreted calculi views them as entities consisting of three components: a set of axioms, a model, and an interpretation function that unambiguously correlates the terms of the axiom system with the entities in the model. In a more general setting, it would be more adequate to say that it is not just one model (and the corresponding interpretation) but a whole class of models (with their corresponding interpretation functions) that belong to the identity of a theory. Normally, any axiom system will have an indefinite number of models. For example, there are many thousands of groups of people, besides the Hendersons, that fulfil the axioms Ax. 1 – Ax. 7 and are therefore families. The same is true of practically any scientific theory. The fact that, normally, we'll have a whole class of models constituting the theory's identity, and not just one, is not a trivial fact and has some important consequences for our understanding of the nature of scientific theories. We'll come back to this point further on.

Have we now solved our original problem of how to determine the meaning of the basic concepts of a theory and of its identity? The answer is: Yes and no. In a purely formal-structural sense, we have: We now know that a theory consists of three components: an axiom system, a model, and an interpretation function, or assignment, connecting the theory's basic terms with reality. But, in a deeper epistemological sense, we haven't solved the problem completely. The reason is that we certainly know there must be such an interpretation function but we still don't know how to construct it. What are the assignments that are really acceptable to provide a reality-based interpretation to the theory's concepts? Or else can we choose them arbitrarily? There are good reasons not to answer affirmatively to the latter question.

Consider the following example. Someone takes the rational numbers 1/3, 2/3 and 1 as TSF's universe of discourse, i.e. the variables in TSF's axioms run only over the set {1/3, 2/3, 1}. (We assume the fractions have been fully simplified.) Further, the term "E" of the axiom system is interpreted as producing a new number from adding to a given number another number, that is "xEy" should mean "there is a number n such that $x+n=y$"; and "M" is interpreted as "having the same denominator as". The reader can check for himself that this interpretation makes TSF's axioms true. Therefore, we should say that the set {1/3, 2/3, 1}, together with the new interpretations of E and M, provides a model for the theory of standard families. {1/3, 2/3, 1} comes out as a family... This is, of course, a very awkward result and it seems to be fully untenable. However, it is not easy to see how this result can be circumvented. Remember that we have argued that any intuitive associations or preconceptions about

UNESCO – EOLSS SAMPLE CHAPTERS

the 'real' meaning of a theory's basic terms should play no role in a scientific context. Their meaning is determined solely by the axioms and the models. But, then, the Hendersons and {1/3, 2/3, 1} are on the same footing; both are families according to our theory.

Some authors have indeed pleaded for this understanding of the semantic structure of scientific theories: Any domain of entities with the appropriate properties that fulfils a theory's axioms should be regarded as a part of the theory's identity. The contrary opinion would be just a remnant of pre-scientific prejudices. However, most philosophers of science reject this position, and for good reasons. The view on the nature of scientific theories we just laid out might be acceptable for purely mathematical theories, but not so for *empirical* theories. The goal of a genuinely empirical theory is to be anchored in the world of experience, and it is this "anchoring" *too* that provides their meaning to the basic concepts. The conceptual framework (and therefore the statements we make within it) has to be *interpreted* in the *empirical* reality. This implies that their models, of whatever kind they might be, must be at any rate *empirical* models, i.e. structured domains that, however idealized and simplified, have something to do with human experience in the last analysis. The problem is how to make sure that the theory's basic concepts, as they are interpreted in its models, get anchored in experience, how they get an empirical meaning without having recourse to vague everyday intuitions.

6. The Radical Empiricist View of Scientific Theories

The most natural move to solve the problem just stated seems to be to say that only those interpretation functions should be admitted that assign some *experiential* items (individual objects, properties, relations) to the variables and basic terms of empirical theories, that is, things that can be the objects of our sensorial experience as human beings, things that we can see, touch, hear, and so on. The 'purest' version of this conception is *phenomenalism*. According to it, the interpretation correlates of the terms of an axiomatized empirical theory consist solely of phenomena as they appear in the flow of sensorial experience to a normal human subject before any theoretical elaboration. An example of a phenomenon in this sense would be "a round red spot in the center of my visual field". These are, in the last analysis, the kinds of things a scientific theory speaks about; phenomena, and only them, build the empirical content of any theory. And when a scientific term doesn't seem, prima facie, to refer to sensorial phenomena, then we should take care of defining it rigorously, by means of more or less lengthy and complicated definition chains, in terms that directly and obviously refer to phenomena. And if the term in question is not definable in this way, it (and the theory containing it) should be dismissed as speculative or 'metaphysical'.

The most prominent example of an attempt to reconstruct the empirical content of science in purely phenomenalistic terms is provided by Rudolf Carnap's great work *Der logische Aufbau der Welt ("The Logical Structure of the World")*, published in 1928. There, starting with a minimal conceptual basis of a purely phenomenalistic nature (actually just one phenomenal relation), Carnaps tries to define rigorously all essential notions of psychology, physics, and the social sciences. In spite of Carnap's impressive constructions, the phenomenalistic program was soon abandoned by most philosophers

UNESCO – EOLSS SAMPLE CHAPTERS

of science, including Carnap himself. The main reason was that it appeared to be confronted with insurmountable difficulties. In particular, the basic concepts of physics appeared to be, as a matter of principle, not amenable to a definition solely in terms of phenomena.

Subsequently, a slightly more moderate empiricist explication of the meaning of scientific terms was proposed by Carnap, Otto Neurath, and others: so-called "*physicalism*". According to it, the meaning of the basic concepts of empirical theories should be given by middle-sized physical objects (and their properties and relationships) as they can be directly observed by a normal human being – things like trees, tables, cables, and so on, and their properties and relations like being tall, or lying upon something, etc. These kinds of items constitute the ultimate empirical content of any genuinely scientific theory. Certainly, physicalists were well-aware that more sophisticated scientific concepts like "electron", "black hole", or "tectonic plate", which occur in renowned scientific theories, look quite different from the concepts of a tree or a cable. Whatever their interpretation correlates might be, they are certainly not things that we can directly observe. But they contended that, with enough patience and technical skills, it should be possible, again, to define them rigorously by long definition chains in terms of the simple concepts referring to what is immediately observable. Otherwise, science would contain irremediably obscure, 'metaphysical' elements, not controllable by experience.

Unfortunately, the physicalistic program was also soon to confront very serious difficulties which made it highly implausible. Careful logical analysis showed that it was not possible, in general, to define the more sophisticated concepts of advanced scientific theories out of the plain notions of the physicalistic basis. The general reason for this situation is that sophisticated scientific concepts prove to have essentially more content that all that can be given by observation – they have a 'surplus meaning'. This fact led to the development of a more refined conception of the nature and semantic structure of scientific theories, which is still basically empiricist in spirit but in a less stringent way than its radical empiricist forerunners. This conception can be said to constitute the 'core' of classical philosophy of science – sometimes also called the "received view" about empirical science. To its examination we devote the next section.

-
-
-

TO ACCESS ALL THE **48 PAGES** OF THIS CHAPTER,
Visit: http://www.eolss.net/Eolss-sampleAllChapter.aspx

Bibliography

Balzer, W., Moulines, C.U., Sneed, J.D. (1987). *An Architectonic for Science*. Dordrecht: Reidel

UNESCO – EOLSS SAMPLE CHAPTERS

(Kluwer). [The standard, most comprehensive exposition of metatheoretical structuralism].

Balzer, W., Sneed, J.D., Moulines, C.U. (eds.) (1998). *Structuralist Representation of Knowledge*. Amsterdam: Rodopi. [An anthology of structuralistic reconstructions of a number of particular scientific theories from different disciplines of the natural and social sciences].

Braithwaite, R.B. (1953). *Scientific Explanation*. Cambridge : Cambridge University Press. [A classical exposition of the two-levels view of scientific theories, and in particular of the Ramsey-method].

Carnap, R. (1928). *Der logische Aufbau der Welt*. Berlin: Weltkreis. [The most outstanding attempt at a phenomenalistic reconstruction of science].

Carnap, R. (1952). *The Continuum of Inductive Methods*. Chicago: University of Chicago Press. [Carnap's classical work on inductivism].

Carnap, R. (1966). *Philosophical Foundations of Physics*. New York: Basic Books. [It contains, among other topics of the philosophy of science, Carnap's latest exposition of the two-levels view of scientific theories, and in particular of the Ramsey-method].

Diez, J.A. and Moulines, C.U. (1997; revised edition 1999). *Fundamentos de Filosofia de la Ciencia*. Barcelona: Ariel. [It includes, among other topics in the philosophy of science, a systematic exposition and comparison of the different approaches on the structure of scientific theories].

Hanson, N. R. (1958). *Patterns of Discovery*. Cambridge: Cambridge University Press. [It argues at length for the thesis of the theory-ladenness of observation].

Henkin, L., Suppes, P., Tarski, A. (eds.) (1959). *The Axiomatic Method*. Amsterdam: North-Holland. [A very comprehensive exposition of the axiomatic method both in general and in its application to the reconstruction of particular scientific theories; it contains, among others, a contribution by E. W. Adams in which this author explains his modelistic view and in particular the notion of "intended interpretations"].

Nagel, E. (1961). *The Structure of Science*. New York: Harcourt. [The most complete presentation of the classical philosophy of science, and in particular of the two-levels view of theories].

Popper. K. R. (1958). *The Logic of Scientific Discovery*. London: Hutchinson. [The classical presentation of falsificationism].

Sneed, J.D. (1971, revised edition 1979). *The Logical Structure of Mathematical Physics*. Dordrecht: Reidel (Kluwer). [The pioneer work in metatheoretical structuralism].

Stegmüller, W. (1973). *Theorienstrukturen und Theoriendynamik*. Berlin-Heidelberg: Springer [An introductory work on metatheoretical structuralism].

Suppes, P. (1970). *Set-theoretical Structures in Science*. Stanford: Stanford University Press. [A very didactic exposition of the set-theoretical method of reconstructing theories with several examples. Unfortunately, the book appeared only as a mimeographed copy that is quite difficult to obtain].

Suppes, P. (1993). *Models and Methods in the Philosophy of Science*. Dordrecht: Kluwer. [It includes a number of essays on the method of set-theoretical reconstruction of scientific theories and on representationalism].

Van Fraassen, B. (1980). *The Scientific Image*. Oxford: Clarendon Press. [The standard exposition of constructive empiricism].

UNESCO – EOLSS SAMPLE CHAPTERS

Biographical Sketch

C. Ulises Moulines , born in Caracas, 1946. MA in Philosophy from the University of Barcelona,1971. PhD in Philosophy of Science from the University of Munich, 1975.

He has been Tenured Professor at the National University of Mexico, the University of Bielefeld (Germany) and the Free University of Berlin; currently, he is the Director of the Institute for Philosophy, Logic, and Theory of Science at the University of Munich. Among other books, he has published: *An Architectonic for Science* (co-authored with W. Balzer and J. D. Sneed), Dordrecht: Reidel (Kluwer), 1987; *Fundamentos de Filosofia de la Ciencia* (co-authored with J. A. Diez), Barcelona: Ariel, 1997/1999; and *La philosophie des sciences – L'invention d'une discipline*, Paris: Editions de la Rue d'Ulm, 2006. His main research areas have been and continue to be the general philosophy of science, the philosophy of physics, the theory of knowledge, and the history of analytic philosophy.

Prof. Dr. Moulines is Member of the Bavarian Academy of Science, of the *Conseil d'Administration* of the *Ecole Normale Supérieure* in Paris, of the International Academy for the Philosophy of Science, and of the German Society for Analytic Philosophy; he is also Honorary Member of the Spanish Society for Logic, Methodology and Philosophy of Science, and of the Catalan Society for Philosophy.

UNESCO – EOLSS
SAMPLE CHAPTERS